AD-A072 134 FLORIDA STATE UNIV TALLAHASSEE DEPT OF STATISTICS F/G 12/1

ON RATES OF CONVERGENCE IN THE L2 NORM OF NONPARAMETRIC PROBABI--ETC(U)

JUL 79 K F CHENG, R J SERFLING N00014-76-C-0608

UNCLASSIFIED FSU-STATISTICS-M506 NL

1 OF 1

AD

A072134

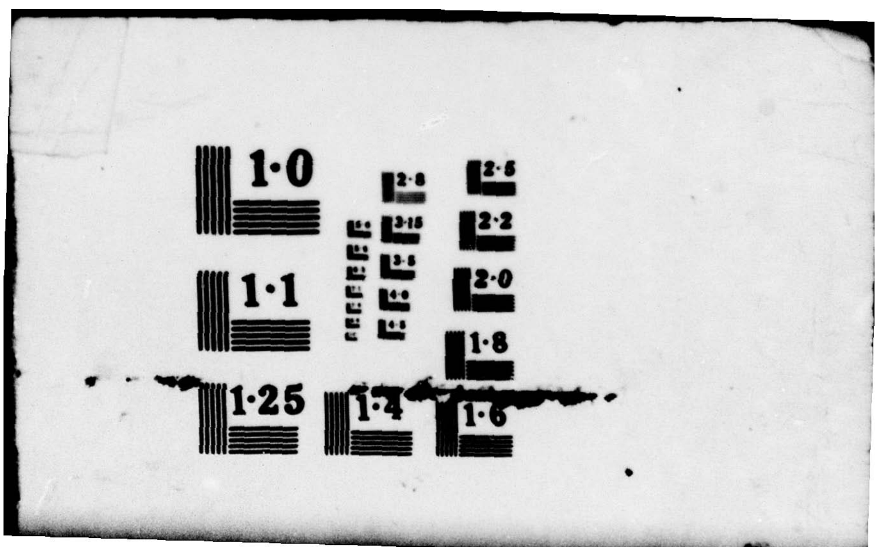
END

DATE

FILMED

9 79

DDC


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5

he Florida State University

(11)

(6)

ON RATES OF CONVERGENCE IN THE L_2-NORM OF NONPARAMETRIC PROBABILITY DENSITY ESTIMATES,

(10) ~~by~~ K. F. Cheng ~~and~~ R. J. Serfling

~~Statistics Report M506~~

~~ONR~~ (9) Technical Report, ~~No. 139~~

(14)

FSU-STATISTICS-M506, TR-139-ONR

(12) 15 p.

(11) July ~~, 1979~~
~~Department of Statistics~~
The Florida State University
Tallahassee, Florida 32306

(15)
Research supported by the Army, ~~Navy and Air Force~~ under Office of Naval Research Contract No. N00014-76-C-0608. Reproduction in whole or in part is permitted for any ~~purpose of the~~ United States Government.

400 277

ABSTRACT

ON RATES OF CONVERGENCE IN THE L_2-NORM OF NONPARAMETRIC PROBABILITY DENSITY ESTIMATES

For estimation of a probability density function f by an empirical probability density function f_n based on a sample of size n from f, a useful measure of distance is the L_2-norm $||f_n - f|| = (\int[f_n(x) - f(x)]^2dx)^{\frac{1}{2}}$. Considerable study of the rate of *mean square* convergence of $||f_n - f||$ to 0 has taken place. This paper investigates the rate of *almost sure* convergence of $||f_n - f||$ to 0, and characterizes moments $E\{||f_n - f||^k\}$ of order $k > 2$ as well. Application to certain estimation problems in nonprarmetric inference is discussed.

Accession For	
NTIS GRA&I	☑
DDC TAB	☐
Unannounced	☐
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or special
A	

Key Phrases: Convergence in L_2-norm; Nonparametric density estimation.

1. Introduction. Consider estimation of a probability density function f, defined on the real line, by an empirical probability density function f_n based on a random sample of size n from f. A natural, useful and interesting measure of the distance between f_n and f is given by $||f_n - f|| = (\int[f_n(x) - f(x)]^2 dx)^{\frac{1}{2}}$, $||\cdot||$ denoting the norm of the space $L_2(-\infty, \infty)$. Indeed, considerable attention in the literature (see review papers by Wegman (1972a,b) and Fryer (1977)) has been devoted to the rate of *mean square* convergence of $||f_n - f||$ to 0, that is, the rate of convergence of $E\{||f_n - f||^2\}$ to 0. Further, the *almost sure* (a.s.) convergence of $||f_n - f||$ to 0 has been studied, by Nadaraya (1973), but without investigating the *rate*. This latter question is explored in the present paper. Also, higher moments $E\{||f_n - f||^{2r}\}$ are characterized.

Besides having intrinsic interest, the results have application (see K. F. Cheng and Serfling (1979)) in connection with estimation of efficacy-related parameters such as $T(f) = \int f^2(x)dx$ arising in certain nonparametric inference problems. For this and other such $T(f)$, the approximation of the estimation error $T(f_n) - T(f)$ by the Gâteaux derivative of $T(\cdot)$ at f with increment $f_n - f$ leads to an error term proportional to $||f_n - f||^2$. In order, then, to obtain useful characterizations of the behavior of $T(f_n) - T(f)$, it is necessary that $||f_n - f||^2$ satisfy appropriate conditions. For example, in order ultimately to derive the central limit theorem for $T(f_n) - T(f)$, we require

$$\text{(A)} \qquad n^{\frac{1}{2}}||f_n - f||^2 \to_p 0.$$

To obtain an associated Berry-Esséen rate of order $O(a_n)$, we require a somewhat stronger result of the form

(B) $$P(n^{\frac{1}{2}}||f_n - f||^2 > a_n) = O(a_n),$$

where $a_n \to 0$. Finally, to characterize the a.s. order of magnitude of the fluctations of $T(f_n) - T(f)$, we require

(C) $$n^{\frac{1}{2}}||f_n - f||^2 =_{a.s.} o(g(n)),$$

where $g(n)$ is a function, such as $(\log n)^{\frac{1}{2}}$ or $(\log\log n)^{\frac{1}{2}}$, tending to ∞.

We shall obtain properties (A), (B), (C) as consequences of bounds on the moments $E\{||f_n - f||^{2r}\}$. Theorems 1 and 3 provide suitable bounds under different types of assumptions on f and f_n, and Theorems 2 and 4 provide corresponding implications germane to (A), (B) and (C). We also obtain a further result apropos to (C) (and (A)) by a different approach, utilizing the connection between the L_2-norm and the sup-norm (Theorem 5).

We confine attention to estimators f_n of the *kernal* or *window* type:

$$f_n(x) = (nc_n)^{-1} \sum_{i=1}^{n} K((x - X_i)/c_n), \quad -\infty < x < \infty,$$

where $K(u)$ is a "kernel" (sometimes a probability density) and $\{c_n\}$ is a "bandwidth" sequence of positive constants tending to 0.

The results of this paper admit in some instances extensions to higher dimensional data and analogues for other types of nonparametric density estimator f_n.

2. <u>Properties of $||f_n - f||$</u>. We shall develop order bounds for the moments $E\{||f_n - f||^{2r}\}$, for r a positive integer, by the following approach. Using the elementary inequality $|a + b|^k \le 2^{k-1}(|a|^k + |b|^k)$, for k a positive integer, along with $||f_n - f|| \le ||f_n - Ef_n|| + ||Ef_n - f||$,

we readily obtain

$$(*) \qquad E\{||f_n - f||^{2r}\} \le 2^{2r-1}(E\{||f_n - Ef_n||^{2r}\} + ||Ef_n - f||^{2r}).$$

The right-most term in (*) is simply the r-th power of the integrated squared bias, $\int[Ef_n(x) - f(x)]^2dx$, for which the behavior is known from studies of the mean square error of $||f_n - f||$. See Lemmas 2 and 3 below. The other term on the right-hand side of (*) is an r-th order analogue of the integrated variance. The following result adequately characterizes the behavior of this term for the present purposes.

LEMMA 1. *Let r be a positive integer. Assume that* $\sup_x|K(x)|$ *and* $\int|K(x)|dx$ *are finite. Then*

$$(1) \qquad E\{||f_n - Ef_n||^{2r}\} = O((nc_n)^{-r}), \ n \to \infty.$$

PROOF. Put $Y_{ni}(x) = c_n^{-1}K((x - X_i)/c_n) - c_n^{-1}E\{K((x - X_i)/c_n)\}$. Then

$$||f_n - Ef_n||^2 = \int[n^{-1}\sum_{i=1}^{n} Y_{n,i}(x)]^2dx$$

and thus the left-hand side of (1) may be expressed as

$$(2) \qquad n^{-2r}\sum_{i_1=1}^{n}\sum_{j_1=1}^{n}\cdots\sum_{i_r=1}^{n}\sum_{j_r=1}^{n} E\{\prod_{k=1}^{r}\int Y_{n,i_k}(x_k)Y_{n,j_k}(x_k)dx_k\}.$$

By the assumptions on K, the expectations in (2) are finite and, therefore, by Fubini's theorem,

$$E\{\prod_{k=1}^{r}\int Y_{n,i_k}(x_k)Y_{n,j_k}(x_k)dx_k\} = \int\cdots\int E\{\prod_{k=1}^{r} Y_{n,i_k}(x_k)Y_{n,j_k}(x_k)\}dx_1\ldots dx_k.$$

By independence of the $Y_{n,i}(x)$'s, $1 \le i \le n$, the expectation in the integrand is 0 except in the case that each index in the list $i_1, j_1, \ldots, i_r, j_r$ appears at least twice. In this case the number of distinct elements in the set $\{i_1, j_1, \ldots, i_r, j_r\}$ is $\le r$. It follows that the number of ways to choose $i_1, j_1, \ldots, i_r, j_r$ such that the expectation in (2) is nonzero is $O(n^r)$. Moreover, these nonzero expectations are uniformly $O(c_n^{-r})$. Hence

$$E\{||f_n - Ef_n||^{2r}\} = O(n^{-2r}n^r c_n^{-r}) = O((nc_n)^{-r}). \quad \square$$

The case $r = 1$ has been derived by Nadaraya (1974) with explicit constant in the $O(\cdot)$ expression. Turning now to $||Ef_n - f||$, we first cite a further result of Nadaraya (1974), namely that

$$||Ef_n - f||^2 = O(c_n^{2s}),\ n \to \infty, \tag{3}$$

for f belonging to W_s, K belonging to H_s, and s an even integer ≥ 2. Here W_s denotes the set of functions $f(x)$ having derivations of s-th order inclusively, $s \ge 2$, $f^{(s)}(x)$ being a bounded continuous $L_2(-\infty, \infty)$ function, and H_s, s an even integer, denotes the class of kernels $K(x)$ satisfying

$$K(x) = K(-x),\ \int K(x)dx = 1,\ \sup_x|K(x)| < \infty, \tag{4a}$$

$$\int x^i K(x)dx = 0,\ 1 \le i \le s - 1, \tag{4b}$$

$$\int x^s K(x)dx \ne 0,\ \int x^s|K(x)|dx < \infty. \tag{4c}$$

The possibility of (3) under milder restrictions on the smoothness of f has been investigated by E. P. Cheng and Serfling (1979). Let W_1 denote the set of functions f *either* having bounded continuous derivative

in $L_2(-\infty, \infty)$, *or* being Lipschitz on $(-\infty, \infty)$ and vanishing off a finite interval. Let H_1 denote the class of kernals $K(x)$ satisfying

$$K \text{ nonnegative}, \int K(x)dx = 1, \int x^2 K(x)dx < \infty. \tag{5}$$

For $f \in W_1$ and $K \in H_1$, (3) holds with $s = 1$. Further, it is shown that $||Ef_n - f||^2 = O(c_n^{(2q+1)/(q+1)})$ if f is $\mathrm{Lip}(-\infty, \infty)$ and satisfies the q-th order tail restriction $\int_{|x|>t} f(x)dx = O(t^{-q})$, $t \to \infty$ (implied by $\int |x|^q f(x)dx < \infty$), and if K belongs to H_1. In other words, (3) holds with s given by $(q + \frac{1}{2})/(q + 1)$. Therefore, for $s \in (\frac{1}{2}, 1)$, we define W_s to consist of functions f which are Lip $(-\infty, \infty)$ and satisfy the q-th order tail restriction for $q = (s - \frac{1}{2})/(1 - s)$. Also, for $s \in (\frac{1}{2}, 1)$, we define H_s to be identical with H_1.

Combining all these results on $||Ef_n - f||$, we have

LEMMA 2. *Let* $s \in (\frac{1}{2}, 1)$ *or be an even integer. Let* $f \in W_s$ *and* $K \in H_s$. *Then* (3) *holds.*

Let us now utilize Lemmas 1 and 2 in conjunction with (*). For $c_n = An^{-1/(2s+1)}$ the bounds provided by these lemmas for the terms in (*) have the same order, $O(n^{-2sr/(2s+1)})$. Therefore, we have

THEOREM 1. *Let* $s \in (\frac{1}{2}, 1)$ *or be an even integer, and assume* $f \in W_s$ *and* $K \in H_s$. *Let* r *be a positive integer. Then*

$$E\{||f_n - f||^{2r}\} = O((nc_n)^{-r}) + O(c_n^{2rs}) \tag{6}$$

and, in particular, for $c_n = An^{-1/(2s+1)}$,

$$E\{||f_n - f||^{2r}\} = O(n^{-2sr/(2s+1)}).$$

For the case s an even integer and $r = 1$, relation (6) has been given by Nadaraya (1974) with explicit constants in the $O(\cdot)$ expressions. Those constants were crucial to his purpose of estimating the optimal constant A to be used in the choice $c_n = An^{-1/(2s+1)}$. For our purposes, however, these constants are irrelevant but it is important that r may be chosen arbitrarily large. By virtue of this feature, we are able to establish

THEOREM 2. *Let* $s \in (\frac{1}{2}, 1)$ *or be an even integer. Assume* $f \in W_s$, $K \in H_s$, *and* $c_n = An^{-1/(2s+1)}$. *Then:*

(i) *For* $\alpha < s/(2s + 1)$,

$$n^{\alpha}||f_n - f|| \to_{a.s.} 0. \tag{7}$$

(ii) *For* a_n *given by* $Bn^{-\beta}$, *with* B *constant and* $\beta < (s - \frac{1}{2})/(2s + 1)$,

$$P(n^{\frac{1}{2}}||f_n - f||^2 > a_n) = O(a_n). \tag{8}$$

PROOF. Let $\epsilon > 0$ be given. Let r be a positive integer. Applying Theorem 1, we have

$$\begin{aligned} P(n^{\alpha}||f_n - f|| > \epsilon) &\le \epsilon^{-r}n^{2\alpha r}E\{||f_n - f||^{2r}\} \\ &= O(n^{r[2\alpha - 2s/(2s+1)]}). \end{aligned} \tag{9}$$

For $\alpha < s/(2s + 1)$, (7) thus follows by the Borel-Cantelli lemma and the fact that r may be chosen arbitrarily large. Now replace ϵ by $a_n^{\frac{1}{2}}$ and α by $\frac{1}{2}$ in (9) and obtain

$$P(n^{\frac{1}{2}}||f_n - f||^2 > a_n) \le a_n^{-r}O(n^{r[\frac{1}{2} - 2s/(2s+1)]}). \tag{10}$$

For $a_n = Bn^{-\beta}$, the right-hand side of (10) is seen to be $O(a_n)$ under the condition that

$$\beta < \frac{r}{r+1}\,\frac{s-\frac{1}{2}}{2s+1}.$$

For $\beta < (s - \frac{1}{2})/(2s + 1)$, this condition is satisfied for a sufficiently large choice of r and thus (8) follows. □

REMARK. Regarding the properties discussed in Section 1, we see from Theorem 2(i) that properties (A) and (C) hold for $f \in W_s$ and $K \in H_s$, where s satisfies $s/(2s+1) > \frac{1}{4}$, or equivalently $s > \frac{1}{2}$. That is, (A) and (C) hold for *all possibilities* of f and K covered by the conditions of the theorem. Further, property (B) holds for a_n of the form $Bn^{-\beta}$ with $\beta < (s - \frac{1}{2})/(2s + 1)$; as $s \to \infty$, β may be increased toward $\frac{1}{2}$, which corresponds to the optimal Berry-Esséen rate in typical circumstances. □

Up to now we have considered the behavior of $||f_n - f||$ under a range of smoothness and/or tail restrictions directly imposed on f (and with K satisfying compatible restrictions). Let us now address the question relative to restrictions on the characteristic function ϕ_f of f, following Parzen (1962) and Watson and Leadbetter (1963). The characteristic function ϕ_f is said to *decrease algebraically of degree* p > 0 if

$$0 < \lim_{|t|\to\infty} |t|^p|\phi_f(t)| < \infty.$$

The compatible estimator f_n is of kernal type and, moreover, must be of "*algebriac form*": K must be square integrable and its Fourier transform ϕ_K must be a bounded even square integrable function. As a parallel to Lemma 2, we have

LEMMA 3. *Let* ϕ_f *decrease algebraically of degree* p > 0, *and let* f_n *be of algebraic form with* K *such that* $\int|t|^{-2p}[1 - \phi_K(t)]^2dt < \infty$ *and with* $c_n = Cn^{-1/2p}$. *Then*

(11) $$||Ef_n - f||^2 = O(n^{-(2p-1)/2p}).$$

In fact, Watson and Leadbetter establish (11) for $E\{||f_n - f||^2\}$ and with an explicit constant in the $O(\cdot)$ expression. Their result yields (11) by Jensen's inequality and Fubini's theorem.

Combining Lemmas 1 and 3, we easily obtain

THEOREM 3. *Let* ϕ_f *decrease algebraically of degree* $p > 0$ *and let* f_n *be of algebraic form with* K *bounded, absolutely integrable, and satisfying* $\int |t|^{-2p}[1 - \phi_K(t)]^2 dt < \infty$ *and with* $c_n = Cn^{-1/2p}$. *Let* r *be a positive integer. Then*

$$E\{||f_n - f||^{2r}\} = O(n^{-r(2p-1)/2p}).$$

The preceding result is an analogue to Theorem 1. A corresponding analogue to Theorem 2, by a similar proof as before, is the following.

THEOREM 4. *Assume the conditions of Theorem* 3. *Then:*

(i) *For* $\alpha < (2p - 1)/4p$, (7) *holds.*

(ii) *For* a_n *given by* $Bn^{-\beta}$, *with* $\beta < (p - 1)/2p$, (8) *holds.*

Regarding properties (A), (B) and (C) of Section 1, comments similar to the Remark following Theorem 2 apply. That is, (A) and (C) hold if $p > \frac{1}{2}$; as $p \to \infty$, (B) holds for β arbitrarily close to $\frac{1}{2}$, the optimal value.

We now augment the assertions of Theorems 2(i) and 4(i) with an additional result utilizing the following lemma from Serfling (1979). Here we denote by $||\cdot||_\infty$ the sup-norm, $||h||_\infty = \sup_x |h(x)|$, and by $||\cdot||_1$ the L_1-norm, $||h||_1 = \int |h(x)|dx$.

LEMMA 4. *Let* f *be a probability density on* $(-\infty, \infty)$ *satisfying* $\int_{|x|>t} f(x)dx = O(t^{-q})$, $\to \infty$ *implied by* $\int |x|^q f(x)dx < \infty$). *Let* g_n *be a sequence of probability density functions satisfying* $||g_n - f||_\infty \to 0$. *Then* $||g_n - f||_1 = O(||g_n - f||_\infty^{q/(q+1)})$. *If, further,* f *vanishes off a finite interval, then* $||g_n - f||_1 = O(||f_n - f||_\infty)$.

Various theorems in the literature provide conditions on f and on the estimator f_n (in some cases considering f_n other than the kernel type), under which the rate of a.s. convergence of $||f_n - f||$ to 0 is characterized: for some $\gamma > 0$,

(12) $$n^\gamma ||f_n - f||_\infty \to_{a.s.} 0.$$

In particular, under minimal smoothness conditions on f and for suitable K, results of Reiss (1975), Silverman (1978) and Winter (1978) *establish that* (12) *holds for* $\gamma < 1/3$. *Using such results in conjunction* with Lemma 4 and the inequality $||f_n - f|| \le ||f_n - f||_1 \cdot ||f_n - f||_\infty$, we obtain corresponding results for the L_2-norm. We thus have

THEOREM 5. *Let* f *and* f_n *satisfy conditions under which* (12) *holds for some* $\gamma > 0$. *Let* f *also satisfy the* q*-th order tail restriction of Lemma* 4, *where* $q = \infty$ *denotes that* f *has bounded support. Then, for* $\alpha < \gamma(2q + 1)/2(q + 1)$, (7) *holds.*

Let us briefly compare the rates provided by Theorems 2(i), 4(i) and 5. In Theorem 5 let us take γ approximately $1/3(\gamma < 1/3)$ and q approximately 1 ($q > 1$). Then (7) with α approximately ¼ follows. In this case the assumptions of Theorem 2 are comparable for $s = 3/4$, yielding (7) with α approximately 3/10. Likewise the assumptions of Theorem 4 are comparable for $p = 2$, yielding (7) with α approximately 3/8. Thus Theorem

4 is more effective in this comparison. Furthermore, Theorems 2 and 4 (and their foundations Theorems 1 and 3) provide additional information on $||f_n - f||$. Of course, the approach based on Lemma 4 can also be exploited to provide bounds on $E\{||f_n - f||^{2r}\}$, via

$$E\{||f_n - f||^{2r}\} \leq D_{f,q}\, E\{||f_n - f||_\infty^{r(2q+1)/(q+1)}\},$$

for a constant $D_{f,q}$ depending only on f and q. This requires knowledge of the moments of $||f_n - f||$, a topic which has received little attention in the literature. Leadbetter (1963) has established $E\{||f_n - f||_\infty^2\} = O(n^{-1+1/p-\epsilon})$ under conditions similar to those of Theorems 3 and 4. However, the higher moments of $||f_n - f||_\infty$ have not been investigated.

REFERENCES

[1] Cheng, E. P. and Serfling, R. J. (1979). Convergence Rates for the Mean Integrated Squared Errors of Some Nonparametric Density Estimators of Recursive δ-Function Type. Florida State University Statistics Report M505 (ONR Technical Report No. 138).

[2] Cheng, K. F. and Serfling, R. J. (1979). On Estimation of a Class of Efficacy-Related Parameters. Florida State University Statistics Report M507 (ONR Technical Report No. 140).

[3] Fryer, M. J. (1977). A review of some non-parametric methods of density estimation. *J. Inst. Maths Applics* 20, 335-354.

[4] Leadbetter, M. R. (1963). On the Non-Parametric Estimation of Probability Densities. Research Triangle Institute Technical Report No. 11 (Doctoral dissertation, University of North Carolina)

[5] Nadaraya, E. A. (1973). On convergence in the L_2-norm of probability density estimates. *Theor. Prob. Applic.* 18, 808-811.

[6] Nadaraya, E. A. (1974). On the integral mean square error of some nonparametric estimates for the density function. *Theor. Prob. Applic.* 19, 133-141.

[7] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33, 1065-1076.

[8] Reiss, R.-D. (1975). Consistency of a certain class of empirical density functions. *Metrika* 22, 189-203.

[9] Serfling, R. J. (1979). A Variation on Scheffé's Theorem, with Application to Nonparametric Density Estimation. Florida State University Statistics Report M502 (ONR Technical Report No. 137).

[10] Silverman, B. W. (1973). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* 6, 177-184.

[11] Watson, G. S. and Leadbetter, M. R. (1963). On the estimation of the probability density, I. *Ann. Math. Statist.* 34, 480-491.

[12] Wegman, E. J. (1972a). Nonparametric probability density estimation: I. A summary of available methods. *Technometrics* 14, 533-546.

[13] Wegman, E. J. (1972b). Nonparametric probability density estimation: II. A comparison of density estimation methods. *J. Statist. Comput. Simul.* 1, 225-245.

[14] Winter, B. B. (1978). Recursive Nonparametric Estimation of Densities and Some Density Functionals. Department of Mathematics, University of Ottowa.

UNCLASSIFIED

Security Classification of this Page

REPORT DOCUMENTATION PAGE

1. REPORT NUMBERS
FSU No. M506
ONR No. 139

2. GOVT. ACCESSION NO.

3. RECIPIENT'S CATALOG NUMBER

4. TITLE
On Rates of Convergence in the L_2-Norm of Nonparametric Probability Density Estimates

5. TYPE OF REPORT & PERIOD COVERED
Technical Report

6. PERFORMING ORG. REPORT NUMBER
FSU Statistics Report M506

7. AUTHOR(s)
R. J. Serfling

8. CONTRACT OR GRANT NUMBER(s)
ONR No. N00014-76-C-0608

9. PERFORMING ORGANIZATION NAME & ADDRESS
The Florida State University
Department of Statistics
Tallahassee, Florida 32306

10. PROGRAM ELEMENT, PROJECT, TASK AREA AND WORK UNIT NOS.

11. CONTROLLING OFFICE NAME & ADDRESS
Office of Naval Research
Statistics & Probability Program
Arlington, Virginia 22217

12. REPORT DATE
July, 1979

13. NUMBER OF PAGES
11

14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)

15. SECURITY CLASS (of this report)
Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)
Approved for public release: distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS
Convergence in L_2-norm; Nonparametric density estimation.

20. ABSTRACT

For estimation of a probability density function f by an empirical probability density function f_n sub n based on a sample of size n from f, a useful measure of distance is the L_2-norm, ~~$||f_n - f|| = (\int[f_n(x) - f(x)]^2dx)^{\frac{1}{2}}$~~. Considerable study of the rate of *mean square* convergence of $||f_n - f||$ to 0 has taken place. This paper investigates the rate of *almost sure* convergence of $||f_n - f||$ to 0, and characterizes moments $E\{||f_n - f||^k\}$ of order k > 2 as well. Application to certain estimation problems in nonparametric inference is discussed.

abs. val. (f sub n - f)